



PHOENIX FOR ADVANCED
LEGAL & ECONOMIC
C E N T E R PUBLIC POLICY STUDIES
www.phoenix-center.org

Written Statement of
George S. Ford, Ph.D.
Chief Economist
Phoenix Center for Advanced Legal & Economic Public Policy
Studies

Before the
House of Representatives
Committee on Energy and Commerce
Subcommittee Telecommunications and the Internet

Hearing on
H.R. ____
A Discussion Draft Addressing Broadband Mapping
and Data Collection

May 17, 2007

Testimony of George S. Ford, Ph.D.

**Chief Economist, Phoenix Center for Advanced Legal & Economic
Public Policy Studies**

**House Committee on Commerce and Energy
Subcommittee on Telecommunications and the Internet**

**Hearing on H.R. _____, A Discussion Draft Addressing
Broadband Mapping and Data Collection**

May 17, 2007

I. Introduction

Mr. Chairman, Ranking Member Upton, and members of the Subcommittee, good morning and thank you for inviting me back to testify before you today. Three weeks ago I had the honor of testifying before this Subcommittee on the most-recent OECD Broadband Rankings report. While there certainly a great deal of controversy surrounding the broadband rankings published by the OECD and ITU, I think we all agree that better data is needed with regard to broadband availability and subscription in this country and abroad. This hearing today is the result of an effort to make that happen.

As a reminder to the Subcommittee, my name is Dr. George S. Ford, and I am the Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies, a non-profit 501(c)(3) organization that studies broad public policy issues related to governance, social and economic conditions, with a particular emphasis publishing academic-quality research on the law and economics of telecommunications

and high-tech industries. Our research agenda is consistently targeted at providing policymakers information about the important role that pro-entry policies must play in the communications industry. We have written over thirty papers on telecommunications policy in the last nine years, many of which have been published in academic journals. Moreover, we make all of our research—as well as rebuttals by those who do not agree with us—available for free at our website, www.phoenix-center.org.

Before beginning my testimony today, I wish to make it clear that the Phoenix Center makes it a policy not to endorse or support any particular piece of federal or state legislation or proposed regulation. Our mission is not to tell policymakers *what* to think about an issue but *how to think* about it. As such, our contributions to communications policy are decidedly more analytical than most, and we refuse to ignore the institutional realities and economic constraints of the communications business.

My comments today are from the perspective of an economist that uses U.S. Government data to research the communications industry. Before and since receiving my Ph.D. in economics from Auburn University in 1994, I have written and published numerous papers on communications markets, almost all of them empirical in nature. I have used in my various studies data made available by the Federal Communications Commission, the Census Bureau, the Bureau of Labor Statistics, the Bureau of Economic Analysis, internal corporate information, and other government and private sources. As such, I welcome and applaud this Subcommittee's efforts to improve the government's data collection abilities, thereby creating the opportunity for the improved empirical analysis of broadband services and policy.

There are a lot of people giving testimony today, so I will get right to the point. With respect to the collection and mapping of broadband availability data, there are a few things that concern me, and I can summarize these as three categories: completeness, comparability, and consistency. The most vital issue, in my opinion, is that the collection and dissemination of the data not be focused solely on generating maps. Maps are pretty, no doubt, and can be informative with respect to availability. In that purpose, they are exceedingly valuable. But to focus solely on maps precludes many public policy innovations that will arise out of the detailed statistical analysis of the underlying data after it is linked with demographic and other geographic information.

II. Completeness

The primary goal of improving our data collection ability is to better measure the availability of broadband services across the United States so that we can make better informed policymaking decisions. For this reason, we need to make sure we collect the *relevant* data, and to make sure this data is complete.

For example, much has been said about mapping geographic availability by five-digit ZIP codes. Under current rules, a carrier must report the availability of service for each ZIP code in which it has at least one high speed subscriber. However, the ZIP code data suffers because it does not provide us a complete picture of broadband availability. Each ZIP code represents approximately 30,000 people or about 12,000 households, yet it only takes one reported broadband connection among those homes for the FCC to consider that ZIP code “served” by that provider. This clearly is inadequate and at best

this ZIP code data can only provide us trend information. Since the ZIP code data now indicates near complete availability, which is widely accepted as being incorrect, it is now time to shrink our unit of measurement to render a more detailed and accurate picture of broadband availability.

For the broadband data collection effort to be complete, we need more granular data collection. In addition, the unit of measurement needs to have a rational basis and be somewhat stable over time. ZIP codes were designed for optimum routes for mail carriers and are subject to change at any time based on changes in letter mail volumes. This is particularly true of ZIP+4, or the nine-digit ZIP codes that are referenced in the Discussion Draft. The only reason for the +4 is to help sort mail—not to assist policymakers in assessing economic development and broadband deployment. I will discuss this issue further below.

In addition, to have this effort be complete, we need to include *all* types of broadband in the collection effort, regardless of technology or provider. Every year, the FCC publishes a small mountain of data and has an entire division devoted to the collection and dissemination of data. The FCC publishes useful data in its semiannual Local Telephone Competition and High Speed Services for Internet Access Reports, the annual Trends in Telephone Service, Telephone Subscribership Report, the Telephone Penetration Report, the Local Operating Company Quality of Service Report, the Reference Book on Rates, Price Indices, and Expenditures for Telephone Service Report,

Universal Service Support Monitoring Reports, International Traffic Data Reports, the biennial the Statistics of Common Carriers, and the Cable Television Rate Surveys.¹ In addition, the FCC collects, maintains and distributes highly-detailed ARMIS (Automated Reporting Management Information System) reports on service quality and network infrastructure for local exchange companies. There is a lot of useful information in these reports for researchers and economists. But you will note that most of that information comes from only one industry segment—the local telephone companies, who have been traditionally regulated. For broadband service, we need *all* providers to participate—regardless of size, geographic location, ownership structure (*i.e.*, public and private entities) and technology deployed—if we are to have a complete picture of how this vital economic infrastructure is developing.

Completeness also dictates that we need information on the different types and speeds of broadband service that are available. Today many scoff at the FCC's definition of 200 kbps as "high speed" service, but when the FCC made that definition in 1999 that service was more than adequate for e-mail and basic Web browsing, the two dominant Internet applications at the time.² The Discussion Draft makes a different choice—it chooses to report 2 Mbps downstream and 1 Mbps upstream. However, as a

¹ For recent releases of several of these reports, see <http://www.fcc.gov/wcb/iatd/stats.html>.

² *Inquiry Concerning the Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion, and Possible Steps to Accelerate Such Deployment Pursuant to Section 706 of the Telecommunications Act of 1996*, Report, 14 FCC 2d 2398 (Feb. 2, 1999) at ¶ 20 ("We have initially chosen 200 kbps because it is enough to provide the most popular forms of broadband—to change web pages as fast as one can flip through the pages of a book and to transmit full-motion video.").

person who spends his time analyzing this data carefully, I believe focusing on a threshold definition simply mires the debate into muckraking. In my view, the definition should be flexible, and cover a range of offerings, including 200 kbps offerings. Indeed, focusing on whether the current definition of 200 Kbs is “too low” or whether we should change to a higher definition misses the point: the purpose of improving our data collection process is not supposed to set an arbitrary point where once achieved we can rest on our laurels; instead, the data collection process is supposed to help inform our decisions in a dynamic process on how we should formulate pro-entry policies to encourage investment and deployment of advanced broadband deployment to all Americans.

Indeed, choosing a specific and singular threshold for what constitutes “high speed” inevitably locks you in to the technology of the day, and likely ignores the wide range of broadband services available that provide important connectivity for consumers and businesses. For example, the proposed 1Mbps upstream will likely exclude all the new deployments of mobile broadband. We should generally expect mobile broadband to be of lower speed than fixed line broadband, but that does not mean we should exclude it from the count since mobile broadband is a very valuable service. Further, from an empirical perspective, having a single threshold, particularly one that is too high, forces researchers to assume that areas that do not meet that threshold have no broadband service at all. Clearly, that would be an error, and therefore biases the results of the statistical testing. In order to maximize its value, the broadband data should include statistics on a range of speeds beginning at the 200 kbps level if only to allow us to continue to use the historical data in a meaningful way.

III. Comparability

Perhaps the greatest defect in much of the broadband data that is collected today, particularly that summarized by the OECD, is the lack of comparability. In the testimony I provided this Subcommittee last month I gave you several examples about how the OECD broadband statistics are essentially comparing apples to oranges, often with absurd results. While lots of data is nice to have, if it is not possible to compare the statistics from one set of data to another intending to measure the same thing, then the data is near worthless.

As a result, I am concerned about the proposal in Section 4 of the Discussion Draft that the broadband map will be composed of data collected by potentially more than fifty different State or local governmental to collect data pursuant to NTIA grants. This collection may (or may not) be an efficient way of obtaining this data, but guidelines should be provided so everyone is collecting something that is sufficient similar for statistical comparison. Otherwise, the data will not provide researchers like me with valuable and useful information. Instead, we run the risk that its only use will simply be as a “feel good” marketing brochure for U.S. Government or particular state or local governments with “rank” envy that may create an incentive to goose their numbers.

In addition, it is not clear to me that codifying international comparisons, as Section 2(b) requires, is a sensible approach. The proposal to collect detailed availability data in this country is not served well by comparing the results to some unspecified data collection effort elsewhere. We want broadband to be available pretty much everywhere

in this country. The meaningful target is, therefore, near 100% availability. Whether or not Iceland has 10% or 100% coverage is irrelevant. I believe our data collection resources should be directed on availability in the United States, which is a big enough task in itself. International comparisons are inevitable regardless of legislation or regulatory mandates, and in my opinion, not very meaningful.

IV. Consistency

From my perspective as a statistician and researcher, what is most troubling about the proposal is its failure to recognize that ZIP codes—even nine-digit ZIP codes—simply cannot be linked sensibly to demographic data. While researchers like me often crudely assign census demographic data to ZIP codes, in doing so we are not able to utilize all of the best demographic information that the Census Bureau collects. Indeed, the Census Bureau has observed “There is no correlation between U.S. Postal Service ZIP Codes and U.S. Census Bureau geography.”³ Naturally, given the dissatisfaction with the high level of aggregation in ZIP code data, narrowing the geographic bounds of the analysis, say from the ZIP to ZIP+4 level, seems like a sensible proposal. However, as far I can tell, there is no ZIP+4 demographic data available from any source. As a result, the Discussion Draft would collect and disseminate information on ZIP+4 code areas and researchers like me would be able to do very little with it. We could make no

³ U.S. Census Bureau, Answers to Frequently Asked Questions about Census Bureau Geography, Maps and Mapping Engines (visited May 15, 2007) (available at: <http://www.census.gov/geo/www/tiger/tigermap.html>).

claims about the relationships between availability and income, race, age, population density, and so forth. This is severely limiting for the empirical analysis of broadband and broadband policy.

Perhaps the overarching issue here is that while maps look nice and are useful in some respects, the most important public policy innovations will not arise from maps but from the analysis of the underlying data itself. Thus, it is critical that the data be collected and/or disseminated in a form that allows it to be easily linked to the best and most detailed demographic data available. Modern mapping software can present the data in a wide range of geographic units, so making very specific decisions on how the data is collected or presented may not be necessary. But, for the sake of research, the underlying data should be made available publicly and at no charge to researchers in a format that can be linked to demographic data.

For example, it may make sense to disseminate the underlying data in a Census Block Group format. Some of the sophisticated cost models for telecommunications plant provide cost estimates at this level of aggregation, and there are obvious opportunities for research by combining those data sets. That said, converting the data to Census Block Groups will still tend to exaggerate availability just like the ZIP code data, but to a far lesser extent, as the target level for population of a Census Block Group is 1,500. Thus, the aggregation problem is reduced by 20-times when moving from ZIP codes to Census Block Groups. Importantly, data *collection* itself need not be tied to any particular geographic unit. My proposal to use Census Block Groups is purely a matter of data dissemination.

In any case, data collected by states and local governments for the mapping project by Section 4 should be collected in the same manner, so it also can be linked directly to the valuable demographic data collected by the Census Bureau. Moreover, it should be required that the data be made available to researchers promptly and in an easily usable, non-proprietary format. Confidential information can be stripped from this data so as to protect company interests yet still provide researchers like me the tools to perform our own analysis. The provision of summary reports by the agency is, in my view, insufficient, because the agency can have mixed motives when making reports on a topic as crucial to our economic future as broadband services. Summary reports likewise rob researchers of the ability to quantify important relationships about broadband and economic activity using high frequency data with significant variation in demographic and economic characteristics. A great deal of information is lost in aggregation.

V. Suggestions and Conclusion

As I said before, the Phoenix Center does not take positions on legislation or lobby. But what we do is write academic quality reports on communications industry topics, so you might say we have a vested interest in ensuring that the data the United States government collects and provides for broadband services are complete and useful for this purpose.

Therefore, while I have raised some concerns about the Discussion Draft, I welcome this Subcommittee's efforts to rationalize the broadband data collection process. Indeed, expressing the political will to collect more data might have a positive

impact in and of itself. For example, it is my understanding that every time the FCC wishes to collect data from industry or change the way it collects data from industry, it must obtain the approval of the Office of Management and Budget pursuant to the Regulatory Flexibility Act of 1995. To obtain OMB approval, the FCC must show that the proposed reporting requirement minimizes burdens on small businesses and entities. Out of concern that these reporting requests will be met with skepticism, the FCC in both Republican and Democratic Administrations seem to have shied away from proposing substantial changes to its data reporting requirements. Moreover, any new information that is requested, like the broadband ZIP code data, must be tailored so as not to burden small business entities.

As a result, while the FCC each year publishes a small mountain of data, that data is a bit of a crazy quilt. This is not the fault of FCC staff, many of whom I know and have worked with and for whom I believe are of very high caliber. But the current process builds in a disincentive not to change or alter the forms that have been used for years. As a result, the FCC still collects information on the cost of artwork that local telephone companies purchase for their headquarters—but it does not collect information on fiber deployed to secondary schools. These failures can lead to regulatory leaps of faith. For example, the FCC has been confronted several times in the last few years with issues relating to competition in the special access market, that is, fiber connections to business buildings. The FCC does not have access to special access network information of competitive providers of these connections. As a result, it must rely upon parties to submit this information, sometimes under seal, in comments or ex parte communications that are, by definition, self-serving.

Let me summarize my suggestions for this effort:

First, all data collected must apply to all industry segments. Given the Regulatory Flexibility Act, I believe that the FCC would need strong support from Congress to make sure that this data is complete and comprehensive. It would seem logical that collecting less data on the artwork that hangs on the walls of telephone companies and putting those resources into mapping broadband availability would be a useful endeavor.

Second, broadband data should be collected and reported based on different “tiers” of bandwidth and by technology. Moreover, the pool of entities required to report such data must be expanded to all broadband providers, regardless of size, location, ownership structure, or technology deployed.

Third, consistency in data collection and presentation is of paramount importance. I am concerned about provisions in the Discussion Draft that would have fifty different entities collecting and compiling information—one for each state—that would then be placed on a “broadband map” of sorts. Consistent data collection and quality permits empirical research, and this system runs the risk that the data collected will be of little use other than as a public relations gimmick. Guidance on form and substance on what is collected and how it is presented should be explicit.

Fourth, broadband data should be collected on a granular basis that facilitates demographic analysis. It should not be enough to know where broadband service is and is not available—we should want to study the demographic reasons *why* service may or may not be available. The ZIP code data that the FCC currently collects, and the ZIP+4

data that the Discussion Draft proposes that the FCC and NTIA collect, do not serve this purpose. In my view, information should be collected on, or at least assignable in a simple way, to Census Block Groups.

Fifth, researchers should have direct access to the raw data, or as near to raw as possible, so we can pursue our own studies, test hypotheses and develop our own conclusions.

In conclusion, we all can agree that the expansion and modernization of our broadband infrastructure is a critical component of the nation's economic growth potential. But to make sound decisions, we must have sound data, sound empirical analysis, and sound interpretation. Many, if not all, of the policy questions you ask yourselves are empirical questions, and empirical questions can only be answered by empirical means. Better data will lead to a more disciplined approach to broadband policy that will render better results and eliminate the waste of resources devoted to quibbling over bad ideas.

Mr. Chairman, thank you again for the invitation to testify today. I would welcome any questions that you may have.